

Computational Linguistics

Course Information

This advanced-level course introduces symbolic and analytic concepts, tools, and methods for working with natural language in computational systems. This course complements the introductory statistical NLP course, Linguistics 539.

Course objectives

We'll use the Chomsky Hierarchy—a hierarchy of complexity of formal languages and the formal mechanisms that are required to characterize them—as a framework to explore concepts and tools that are commonly used in symbolic computational linguistics: regular grammars, as represented by regular expressions, finite-state automata and finite-state transducers; and context-free grammars. This course also introduces programming that is relevant to computational linguistics in three programming languages: Perl, Python and Prolog (for definite clause grammars). We apply these tools to parsing a small range of realistic language data.

Learning outcomes

Successful students in this course will...

1. be able to read and write simple programs in Perl, Python, and Prolog.[†]
2. be able to apply basic concepts, techniques and tools in symbolic computational linguistics.[‡]
3. be able to explain at least one advanced concept in computational linguistics.
4. be well prepared to take *Advanced Computational Linguistics* (Ling 581) and *Statistical Natural Language Processing* (Ling 539).

[†] relates to Linguistics HLT program outcome #1.

[‡] relates to Linguistics HLT program outcomes #2 and #3.

HLT learning outcomes addressed in this course

1. Students will demonstrate programming skills for the workplace.
2. Students can use fundamental algorithms & concepts of Natural Language Processing.
3. Students can use tools & packages typically used in Natural Language Processing.

Prerequisites

There are no formal prerequisites. However, this version of the course is designed for students who have successfully completed LING 529 and 531, and that background will help students gain the most from this course.

Instructor

name Eric Jackson
email ejackson1@arizona.edu
hours Mondays 2:00pm–4:00pm (Arizona time, UTC-7) in person (COMM 114A) and online via Zoom at <https://arizona.zoom.us/j/89529422793> (passcode 736013), and by appointment.

Requirements

Students are expected to actively participate in the course by watching the lecture videos, reading the assigned readings, completing the assigned homework, and engaging with the instructor and other students in the course forum. You're all adults, and you're responsible for your own learning—which is all the more critical in this asynchronous online course.

Lecture videos will be available on the course website (D2L). You are expected to watch all lectures and understand the content. If the content of a lecture is not clear, you are expected to send a question to the instructor by email, meet with the instructor in regular office hours or arrange another time to meet, or post a question for clarification on the course forum.

There will be seven graded homework assignments, eight forum discussions that are graded for completion only, and ungraded assignments. Although attempting or completing the ungraded assignments will not affect your grade, they *will* affect how much of the course content you understand and retain. I have designed all assignments, both graded and ungraded, to aid your learning, and I expect everyone to complete them, to gain the most from the course.

Graded homework assignments will be given via the course website (D2L). Assignments will be given each week and will be due on Tuesdays at noon, Arizona time. Your completed homework will be collected through the course website and must be in PDF format; files submitted in any other format (.doc, .docx, .rtf, .odt, .txt, or any other) will not be accepted. Freely available options to convert files to PDF include Google Docs and LibreOffice; students also have free online access to Adobe Acrobat by signing in with their UArizona NetID.

Forum discussion items will be graded for completion, not for content. There will be one forum prompt per week, for a total of eight posts.

This course will have a final project in lieu of a final exam. Students will prepare a **10-15 minute presentation summarizing an advanced topic selected from the textbook**. Because this is an asynchronous course, each student will record themselves giving this

presentation and post the video online. Students are required to watch and post feedback for at least two other students' presentations in the final week of the course.

The due date for each assignment will be posted with the assignment in D2L. All times will be given in Arizona time (Mountain Standard, GMT-7). Accepting late work would mean that I cannot give timely feedback to the rest of the class on the issues in that assignment. **Except for university-approved reasons listed below, late work will not be accepted.**

Your overall course grade will be calculated based on this weighting of assignments:

	type	number	total
homework assignments		7	65%
forum interaction		8	10%
book chapter report		1	25%
	total		100%

Readings

A draft version of the textbook used in this course is available for free on-line.

Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Jurafsky and Martin, <https://web.stanford.edu/%7Ejurafsky/slp3/>

Additional readings will be available digitally in the course D2L site.

Technology

This course will use Perl 5, Python 3 (I'll be using Python version 3.10), and SWI Prolog. All of these can be downloaded for free, and guidance on how to install these will be provided in the course materials in D2L. I can provide limited support to get these running in a Linux environment, but if you prefer to have them running in a different operating system, you're welcome to do so. Basic tasks that we'll do in Prolog can alternatively be performed in an online environment. (Information on Prolog alternatives will be posted in D2L.)

You are free to install interpreters for these programming languages either directly in your system or indirectly via Docker, a virtual machine, a dual-boot linux system, etc. You just need to be able to run the commands that the homework asks of you.

Student Work Policy and Collaboration

Students are encouraged to discuss problems and general approaches for solutions, but everyone must turn in their own work. You may not submit assignments that are substantially the same as your classmates. You may not submit work which was simply copied and pasted from generative AI. The purpose of assignments is to get **you** to think about the issues, and

having ChatGPT craft an answer for you does not bring any benefit **to you** in terms of your own learning.

Assignments that seem suspiciously similar, or those that seem to have been produced using generative AI, will be forwarded to the Dean of Students office in accordance with the Code of Academic Integrity (linked below). Please be a responsible adult and don't run the risk of losing credit for an assignment by copying, by allowing others to copy from you, *or* by having ChatGPT do your assignment for you.

Generative AI is a tool, just like a calculator is a tool for doing math, or a bicycle is a tool for transportation. In some contexts, being able to use a calculator is an important skill—while in other contexts, like when you're taking a math test to see whether you know basic math facts, solely using a calculator short-changes your education. A bicycle is a tool that allows us to get from one place to another faster and more efficiently than running—but if you're going to be tested in your time for a 5k run, it won't help you to train for running solely by riding a bicycle. We will probably all need to know how to use generative language models for tasks at some point, but having one write your homework or forum posts for this class is not appropriate. Put in the thinking yourself, and reap the mental benefit for yourself.

Schedule

The course covers four major topics, with roughly two weeks per topic, and lectures and assignments for each module. Each week's material becomes available on a Monday, and all assignments for that week are due by the following Monday (in the case of readings and forum posts) or the following Tuesday (in the case of written assignments).

Week: Topic	Dates	Forum (due)	Asmt (due)
1: Foundations	1/10-1/16	1 (1/15)	1 (1/16)
2: Perl: Basics	1/15-1/23	2 (1/22)	2 (1/23)
3: Perl: Quirks	1/22-1/30	3 (1/29)	3 (1/30)
4: Regular languages: Regular expressions	1/29-2/6	4 (2/5)	4 (2/6)
5: Regular languages: Finite state machines	2/5-2/13	5 (2/12)	5 (2/13)
6: Grammars and Context-free languages	2/12-2/20	6 (2/19)	6 (2/20)
7: Grammars/Your term project presentation	2/19-2/27	7 (2/26)	7 (2/27)
7.5: Feedback on at least two others' presentations	2/26-3/1	8 (3/1)	

These dates are subject to change. Authoritative due dates will be listed for each assignment in D2L. Check the D2L course calendar to make sure you don't forget or miss a deadline.

University boilerplate

All of the following items are required by the university to be included on syllabi. If you find something here that is surprising or unexpected, please bring it up with me as soon as possible.

By way of a brief summary:

Disabilities If you have a disability that affects how you will need to do the work in this class, please let me know *within the first week of class*.

Academic Code of Conduct Cheating and plagiarism are not remotely acceptable in any way. You are responsible for knowing whether your own behavior qualifies as plagiarism. Disruptive behavior in class—which here includes audio, video, or text on any of our course websites or by email—is not acceptable. Please be respectful of others.

Sensitive Material This is a university and you are adults. It is possible that we may touch on topics that some students could find sensitive during the semester. Given the focus of this course, this seems unlikely, but I alert you nonetheless.

Health & Wellbeing

The university has a specific site for COVID information: <http://covid19.arizona.edu>. If you are experiencing personal or financial challenges from any health-related issue, let me know as soon as you can if we need to make accommodations, and please stay safe.

The semester ahead may come with ups and downs in both physical and mental health, but there are lots of ways to support yourself. Eat well, get regular exercise, and don't neglect things like self-care, talking with friends and family, or getting a fresh perspective from a supportive group. Stress is a normal part of life and may even motivate you sometimes, but chronic or overwhelming stress can affect your physical and mental health and wellbeing. Pay attention to your personal signs that you're overly stressed, like changes in your mood, appetite, sleep, behavior, or new physical symptoms (aches, pains, etc.) that interfere with school and daily life. If you notice these signs or have questions about helpful resources, I welcome you to talk with me. You can also visit caps.arizona.edu/mental-health for mental health tools and resources.

Mental Health & Wellness Resources

- **Health & Wellness:** Campus Health provides quality medical, mental health, and wellness services for students. Visit health.arizona.edu or call 520-621-9202 (520-570-7898 for help after hours)
- **Mental Health:** Campus Health's Counseling & Psych Services offers a range of mental health support tools and services like self-care strategies, peer support, groups and workshops, and professional mental health services. Visit caps.arizona.edu/mental-health or call CAPS 24/7 at 520-621-3334 to learn more.

- **Crisis Support:**

Suicide & Crisis Lifeline: call 988 Crisis Text Line: text TALK to 741-741 Visit preventsuicide.arizona.edu for more suicide prevention tips and resources

Absence and Class Participation Policy

Attendance in an all-online course is not evaluated like attendance in an in-person course. For this course, attendance will be represented by active reading, completion, and participation in online course activities, including loading/viewing materials and completing activities posted on D2L, OpenClass, our course forum, and any other related websites.

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at: <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

The UA policy regarding absences is that any sincerely held religious belief, observance or practice will be accommodated where reasonable, <http://policy.arizona.edu/human-resources/religious-accommodation-policy>.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: <https://deanofstudents.arizona.edu/absences>

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities.

Students are asked to refrain from disruptive conversations with others in the course, including on asynchronous course platforms. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue inappropriate behavior will be removed from that venue and may be reported to the Dean of Students.

Threatening Behavior Policy

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See <http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, <https://drc.arizona.edu/>) to establish reasonable accommodations.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of

independent effort unless otherwise instructed. **If you use a code snippet that you came up with from discussions with a classmate, that you found online, or even that you got from a large language model, it's important to cite where it came from, whether that source was Sally Classmate, GitHub.com, stackexchange.com, or ChatGPT.**

Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: <http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity>.

The UA Library provides a helpful learning module for students to understand and avoid plagiarism: <https://libguides.library.arizona.edu/info-strategies/plagiarism>

UA Nondiscrimination and Anti-harassment Policy

The University is committed to creating and maintaining an environment free of discrimination; see

<http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.